



IAASSIC

IAASSIC

CARNET DE LABOREATORE

Projet Master

IAASSIC



UNIVERSITE
BRETAGNE
SUD

UNIVERSITE BRETAGNE SUD

2024/2025

Titre du Projet : IA pour la sûreté des systèmes industriels critiques

Membres du groupe : Ali Elkhalfi et Melissa Iberoualene

Encadrants :

- Mawloud Omar (Tuteur principal)
- Caroline Larboulette
- Vincent Segonne

Date de début : 10 octobre 2024

Objectif du Projet :

Ce projet vise à concevoir et à implémenter une solution d'intelligence artificielle pour renforcer la sûreté des systèmes industriels critiques. L'objectif est de surveiller, détecter, et prévenir les anomalies en temps réel afin de garantir une fiabilité accrue dans des environnements industriels où toute défaillance peut avoir des conséquences graves. Ce travail est essentiel pour assurer la résilience des infrastructures critiques en intégrant des technologies avancées d'IA permettant une réponse proactive aux risques.

Objectif du Semestre 1 (Recherche) :

- Analyser l'état de l'art des méthodes d'IA appliquées à la sûreté industrielle.
- Étudier les approches de détection d'anomalies et de prévention des défaillances dans les systèmes critiques.
- Élaborer un rapport présentant la problématique, l'état de l'art et les pistes de recherche. Le rapport inclura une synthèse des approches actuelles, une analyse comparative et des recommandations pour les méthodes à tester au semestre suivant.

Semaine 1 : 14-18 octobre 2024

- Échanges avec l'encadrant :

- Nous avons été officiellement acceptés sur le projet intitulé « IASSIC ». Une première rencontre avec notre tuteur, Monsieur Mawloud, nous a permis de clarifier ses attentes et les objectifs du projet.

- Lors de cette rencontre, Monsieur Mawloud a insisté sur l'importance de bien comprendre le projet dans sa globalité avant de proposer des idées innovantes. Il a souligné qu'il est essentiel de se familiariser avec les solutions déjà existantes dans le domaine en étudiant comment les travaux précédents ont abordé les problématiques de sûreté et quelles solutions ont été mises en place. Cette démarche nous aidera à identifier les défis actuels et à poser une base solide pour envisager des améliorations ou innovations pertinentes.

- Prochaines Étapes (TODO) :

- Rechercher des travaux antérieurs et des approches similaires en sûreté des systèmes critiques.

- Planifier une nouvelle réunion pour finaliser l'analyse commune des références.

- *Références en IA pour systèmes critiques industriels et de transport:* Notre lecture de l'article *Artificial Intelligence for Safety-Critical Systems in Industrial and Transportation Domains: A Survey servira de base pour identifier les méthodologies spécifiques à ces secteurs.

- *Normes de sûreté :* Explorer les standards, comme ceux évoqués dans l'article de Zhu et al., afin de comprendre comment les normes existantes (ISO, IEC) peuvent encadrer l'intégration de l'IA dans des systèmes sûrs.

- *Confiance et transparence dans l'IA :* Analyser des articles sur l'IA digne de confiance, comme *Trustworthy AI: Industry-Guided Tooling of the Methods*, pour intégrer des critères de transparence et d'interprétabilité dans notre solution.

Semaine 2 : 21-25 octobre 2024

- Travail Réalisé (DONE) :

- Individuel : J'ai analysé l'article intitulé "Artificial Intelligence for Safety-Critical Systems in Industrial and Transportation Domains: A Survey". J'ai identifié plusieurs points clés, notamment les méthodologies utilisées pour assurer la sûreté dans des systèmes critiques et les principaux défis rencontrés pour garantir la fiabilité de l'IA dans ces environnements.

- Binôme : Mon binôme s'est chargé d'analyser deux autres articles, afin de répartir le travail et de couvrir une plus grande variété de perspectives et d'approches. Après nos analyses respectives, nous avons échangé brièvement sur les idées principales dégagées, en vue de préparer une synthèse commune.

- Prochaines Étapes (TODO) :

- Préparer un document de synthèse, en regroupant les points importants des articles analysés, avec un focus sur les méthodologies les plus pertinentes et les limites identifiées pour la sûreté des systèmes critiques. Ce document servira de base pour l'état de l'art du projet.

- Organiser une réunion pour partager en détail nos analyses respectives, comparer les conclusions principales, et finaliser ensemble les éléments à inclure dans la synthèse.

- Revoir et ajuster les tâches en fonction des résultats de la réunion pour s'assurer que tous les points clés sont bien couverts et comprendre les perspectives différentes apportées par chaque article.

- Réflexions et Commentaires :

- La répartition du travail s'est révélée efficace, car elle nous a permis de couvrir plus de terrain en peu de temps. Ce partage des perspectives a enrichi notre compréhension des méthodologies et des défis de sûreté dans les systèmes critiques. Une réunion dédiée nous aidera à harmoniser nos points de vue et à organiser nos idées de manière cohérente pour l'état de l'art.

Semaine 3 : 28 octobre - 1er novembre 2024

- Travail Réalisé (DONE) :

- *Synthèse des articles* : Nous avons préparé un document de synthèse en regroupant les points importants des articles analysés, en nous concentrant sur les méthodologies pertinentes et les limites rencontrées dans les approches existantes pour la sûreté des systèmes critiques.

- Problématiques explorées:

- Cette semaine, nous avons également élargi notre recherche pour identifier d'autres problématiques rencontrées dans le domaine de la sûreté des systèmes critiques, afin d'évaluer si des solutions innovantes pourraient y être appliquées. Nous avons listé plusieurs défis auxquels les approches existantes ne répondent pas entièrement et envisagé des pistes potentielles à explorer.

- Prochaines Étapes (TODO) :

- Réunion de binôme : Planifier une nouvelle réunion pour discuter de l'ensemble des problématiques identifiées et des solutions envisageables, afin de bien comprendre chaque point avant de rencontrer notre tuteur.

- Rendez-vous avec le tuteur : Préparer un rendez-vous avec Monsieur Mawloud pour lui présenter notre avancement, partager nos découvertes et obtenir ses retours sur les pistes de solutions envisagées. Ce rendez-vous nous permettra également de clarifier ses attentes pour les prochaines étapes et les démarches à entreprendre. `

- Réflexions et Commentaires :

- La réflexion sur des problématiques plus vastes nous a permis de mieux cerner les lacunes des solutions actuelles et d'envisager des améliorations innovantes.
- Exploration des "Safety Bags" : L'idée des "safety bags" pour surveiller les décisions de l'IA en temps réel, abordée dans l'article *Artificial Intelligence for Safety-Critical Systems*, est intéressante. Intégrer un mécanisme de ce type pourrait renforcer la fiabilité de notre solution en assurant qu'aucune prédiction dangereuse ne soit mise en œuvre sans validation préalable.
- La prochaine réunion avec le tuteur sera l'occasion de valider ces pistes et de voir si d'autres solutions similaires peuvent être envisagées pour combler ces lacunes.

Lecture et résumé des articles proposé :

Here are some key points and a summarized analysis based on the reference document "Artificial Intelligence for Safety-Critical Systems in Industrial and Transportation Domains: A Survey" that would be helpful for your lab notebook.

Evaluation of the Source:

The report provides a thorough analysis of the use of AI in safety-critical systems, with a focus on the transportation and industrial sectors. It highlights the particular difficulties in incorporating AI in these fields, where dependability and adherence to safety regulations are crucial.

Important Points:

1. AI and Functional Safety Challenges:

- **Explainability and Analyzability** : AI models, which are sometimes viewed as "black boxes," are opaque, making it difficult to understand their judgments, particularly in situations with significant stakes.

- **Standards and Compliance** : Because AI is probabilistic, it is challenging to align traditional safety engineering standards like IEC 61508 or ISO 26262, necessitating the development of innovative methods to satisfy compliance requirements.

2. Types of AI in Safety Applications:

- Machine Learning (ML): Widely used for tasks like anomaly detection, but its stochastic nature creates unpredictability.

- Connectionist Networks (e.g., Deep Learning): Highly valued for complex perception tasks but pose significant verification challenges.

- Symbolic AI and Optimization: Provide logical reasoning and deterministic outcomes, beneficial for systems requiring high transparency and provability.

3. Safety Assurance Techniques:

- Safety Bags and Safety Envelopes: In situations when formal verification is difficult or impracticable, techniques such as runtime monitors (safety bags) help guarantee that AI decisions do not result in dangerous situations.
- Although scalability problems occur with complicated models, formal verification and model checking are crucial for validating important system components.

4. Trustworthiness and Compliance Dimensions:

- Trust in AI: Trust goes beyond engineering to include moral and legal considerations, which are essential for social acceptance.
- AI Lifecycle Compliance: Although it is still a work in progress, it is essential to make sure AI complies with standards at every level, from data management to deployment.

5. Future Directions:

- Continued research is needed in areas like explainability, provability, and robustness to enhance AI's role in safety-critical domains.
- The document suggests that AI standards, particularly for autonomous systems, are evolving but require robust frameworks to assure long-term safety and reliability.

Understanding the larger context of AI in safety-critical systems will be made easier with the help of this analysis. Including these ideas in your lab notebook for your “next steps” could involve:

Outlining certain methods applied in each form of AI, Emphasizing the importance of explainability and reliability in the goals of your project; scheduling sessions with your binôme to go over how each AI safety strategy might apply to your particular system configuration.

Sources et Références :

1. Boeing 737 MAX Crashes Explained – Les raisons des accidents du Boeing 737 MAX et les mesures de sécurité prises pour remédier aux problèmes.
2. USS Fitzgerald et USS John S. McCain – Rapport sur les collisions impliquant ces navires et les analyses sur les causes de ces incidents.
3. Artificial Intelligence for Safety-Critical Systems in Industrial and Transportation Domains: A Survey
4. Trustworthy AI: Industry-Guided Tooling of the Methods